Платформа для любой биометрии: проблемы и решения, Или как устроена Единая биометрическая система.

Сергей Браун

Артур Душелюбов











Платформа для любой биометрии: проблемы и решения, Или как устроена Единая биометрическая система.



Сергей Браун

Заместитель директора
департамента цифровой
идентичности

E-mail: sergey.braun@rtlabs.ru

Tel.: +7 (919) 784 – 89 – 50



Артур Душелюбов

Начальник отдела развития
и разработки департамента цифровой
идентичности

E-mail: artur.dushelyubov@rtlabs.ru

Tel.: +7 (926) 670 — 52 — 06

Эта презентация без комментариев докладчика не может считаться законченным материалом и отдельное рассмотрение может привести к неправильным трактовкам







Единая биометрическая система (ЕБС)



236
банка



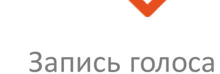
>13 500
отделений



1048
городов

ЕБС – это национальная платформа, которая предоставляет возможность гражданам пользоваться государственными и коммерческими услугами удобно и безопасно.

Биометрические параметры




Финансовый рынок





Открытие вкладов


Получение кредитов

Коммерческие и государственные услуги


Государство
Предоставление услуг по биометрии


Нотариат
Предоставление услуг по биометрии


Ритейл
Оплата покупок одним взглядом


E-commerce
Оплата онлайн по биометрии с подтверждением возраста


Образование
Удаленная сдача экзаменов


Суды
Участие в судебных заседаниях по видеосвязи


Транспорт
Проход/оплата проезда с использованием биометрии


Телеком
Дистанционное получение e-SIM







Работа с биометрией – ситуация…

ПРОБЛЕМА

Работа с лицом и голосом:

- Существует много решений (вендоров)
- Различные характеристики биопроцессоров
- Нейросети постоянно развиваются

Безопасность:

- Проверить, не фото ли перед нами?
- А не запись ли это?
- Все хотят оценить риски ошибки

РЕШЕНИЕ

- Проверять по нескольким модальностям с несколькими сетями – хорошо!
- Отбить атаки – необходимо…
- И всё это без потерь данных, отказоустойчиво, под нагрузкой…







Нейросети и попугаи…

Сеть 1 (сравнение) → 42

Сеть 2 (сравнение) → 98

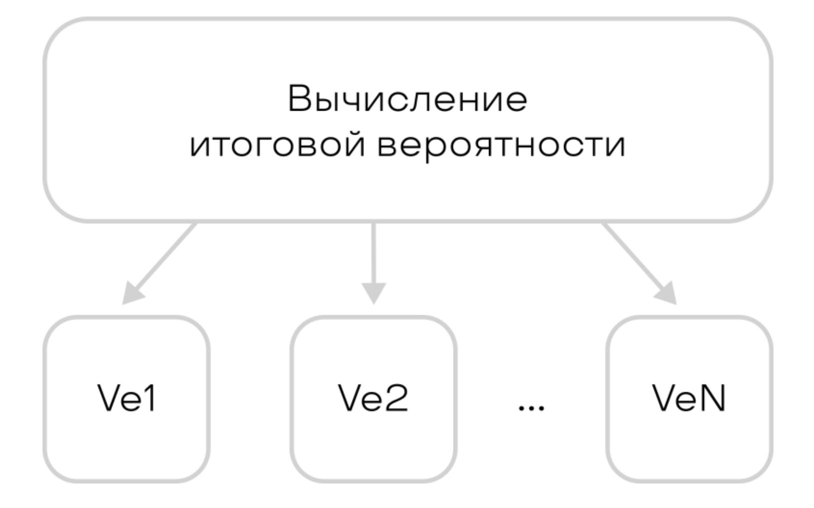








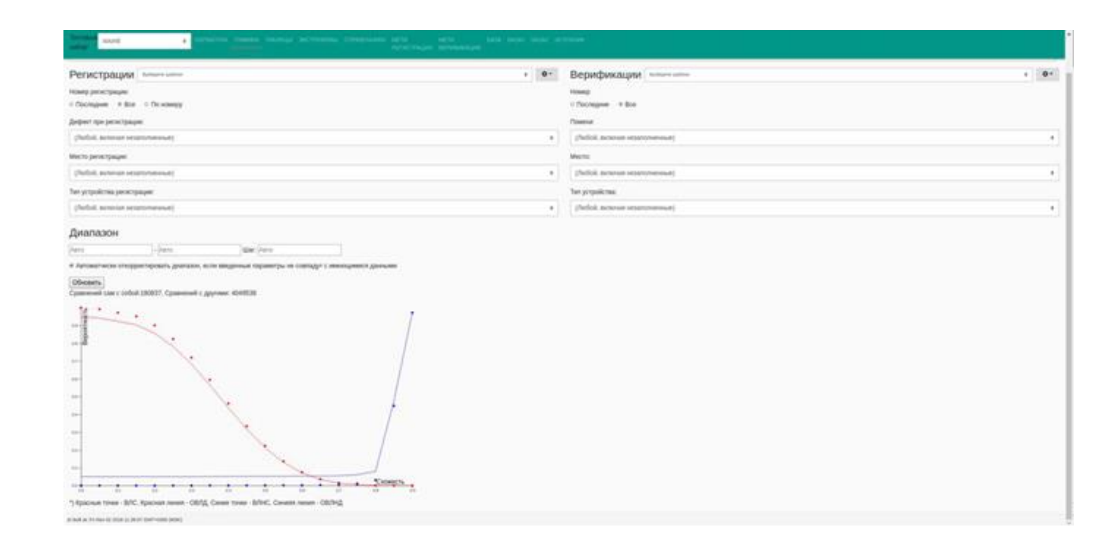
Вендоры возвращают score



Собрали базу образцов…

Измеряем:

- Вероятность ложного допуска
- Вероятность ложного недопуска
- Вероятности ложного несовпадения
- Вероятности ложного совпадения
- Обобщенную вероятность ложного допуска
- Обобщенную вероятность ложного недопуска









Этап 1

Измерение показателей нейросетей







Нейросети и вендоры

Свойства нейросетей:

- Вендоры меняют нейросети
- Все вендоры не всегда извлекают шаблон
- У всех различное API

Что мы сделали:

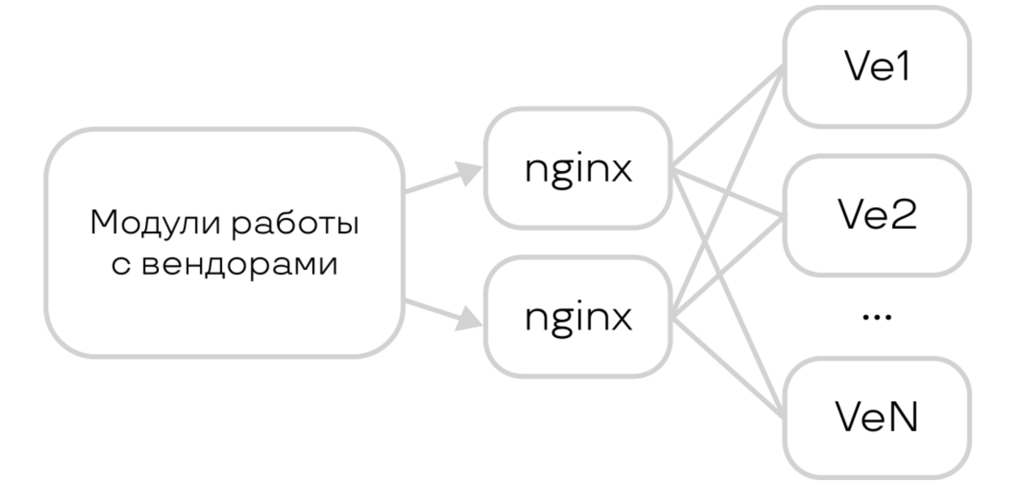
- Делаем переиндексацию.
- Регистрируем образцы во всех биометрических процессорах
- Делаем общее API (extract/compare/…)







Шаблоны и смена вендора



Порядок регистрации:

1. Происходит считывание биометрических образцов.
2. Из конфигурации берётся список вендоров.
3. По каждому из вендоров отправляется запрос на извлечение БШ.
4. RR для распределения нагрузки (модуль)
5. Происходит выбор конкретного биометрического процессора вендора (nginx).
6. При получении ответов от всех вендоров (либо по истечении тайм-аута), извлеченные биометрические шаблоны отправляются на сохранение.

Подключение нового вендора:

- Проводим тестирование и измерения
- Отправляем на биопроцессор все образцы по модальности
- Сохраняем результаты
- Активируем вендора в системе!







API для биопроцессоров

```
POST /pattern/v1/extract HTTP/1.1
Content-Type: multipart/form-data;
boundary=f3URHA_Xnhk0D8gW1iCGLPQk9_gjZr_ywsH
Content-Length: {длина тела сообщения}

--f3URHA_Xnhk0D8gW1iCGLPQk9_gjZr_ywsH
Content-Disposition: form-data; name="sample"
Content-Type: audio/pcm

{Поток байт биометрического образца}
--f3URHA_Xnhk0D8gW1iCGLPQk9_gjZr_ywsH

Content-Disposition: form-data; name="metadata"
Content-Type: application/json
{
 "voice_fragments": {
   "digits_asc": {
     "start": 1.001,
     "end": 9.111
   },
   "digits_desc": {
    "start": 10.999,
     "end": 22.504
   },
   "digits_random": {
     "start": 24.001,
     "end": 32.001
   }
 }
}
--f3URHA_Xnhk0D8gW1iCGLPQk9_gjZr_ywsH
```

```
POST /v1/pattern/compare HTTP/1.1
Content-Type: multipart/form-data;
boundary=f3URHA_Xnhk0D8gW1iCGLPQk9_gjZr_ywsH
Content-Length: {длина тела сообщения}

--f3URHA_Xnhk0D8gW1iCGLPQk9_gjZr_ywsH
Content-Disposition: form-data; name="bio_feature"
Content-Type: application/octet-stream

{Поток байт шаблона, полученный из биометрического образца при
биометрической верификации}
--f3URHA_Xnhk0D8gW1iCGLPQk9_gjZr_ywsH

Content-Disposition: form-data; name="bio_template"
Content-Type: application/octet-stream

{Поток байт Биометрического контрольного шаблона}
--f3URHA_Xnhk0D8gW1iCGLPQk9_gjZr_ywsH--
```







Этап 2

Работа с биометрией

Измерение показателей нейросетей

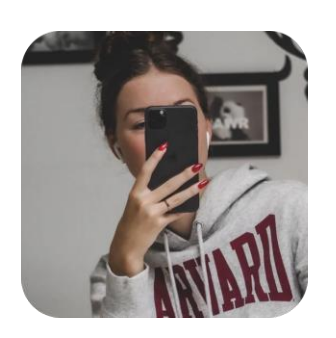




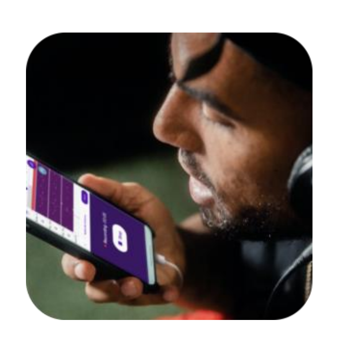


Как обезопасить себя и клиента?

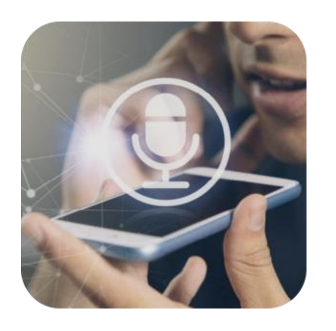
Возможности Лайвнесс:



Проверки
по фото

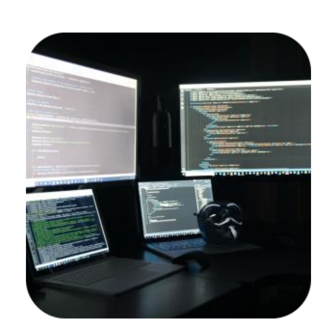


Проверки
по голосу

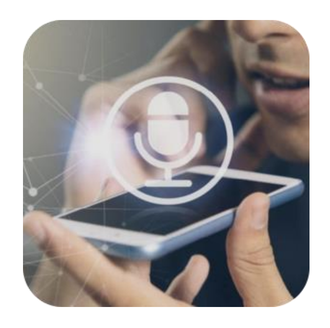


Комбинированные
проверки

Решение:



Использовать всё
что можно!



Работаем аналогично
лицу/голосу







Этап 3

Работа с Liveness

Работа с биометрией

Измерение показателей нейросетей







Биометрия, какая ты?

- Нужно использовать видео ~ 20Mb
- Фотографии и звук по 2-3Mb
- Векторы нейронных сетей 1Kb-3Mb
- Всё это доставить в несколько точек

Транспорт

Работа с Liveness

Работа с биометрией

Измерение показателей нейросетей





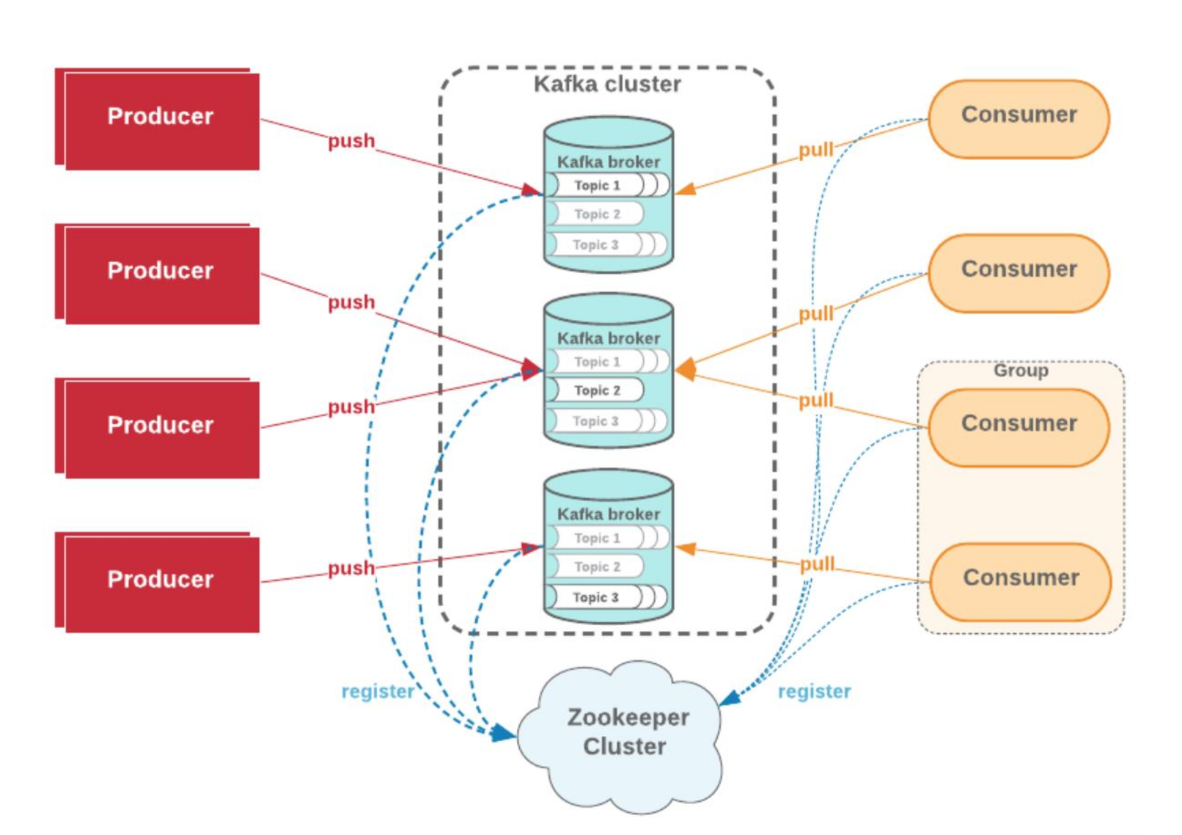


Транспорт и очередь

1. Очереди с персистентностью + балансировка



Apache Kafka architecture







Транспорт и очередь

2. Очереди с репликацией



Репликация

Broker 1	Broker 2	Broker 3
Topic Partition 0 (leader)	Topic 2 Partition 0 (ISR)	Topic 2 Partition 0 (ISR)
Topic 2 Partition 1 (ISR)	Topic 2 Partition 1 (Leader)	Topic 2 Partition 1 (ISR)







Транспорт и очередь

3. Сравнение с RMQ



Feature	Kafka	RabbitMQ
Brokerless/Decentralized	+	–
Persistence/Reliability	+	+
Message Retention	Policy	Acks
Capacity	Cluster Disks	Central Store

Транспорт и очередь

4. Схема с kafka streams для конвейеров обработки

APACHE kafka

Выход

Вход

resp 3

KSApp SaveBio

resp 1

BioProxy

KSApp Verreg

resp 2

Liveness







Этап 4

Работа с Liveness

Работа с биометрией

Транспорт

Измерение показателей нейросетей

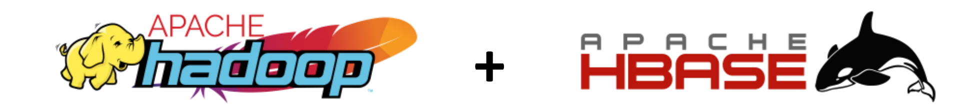






Передавать научились – давайте сохраним!

1. Логика RegionServer'a





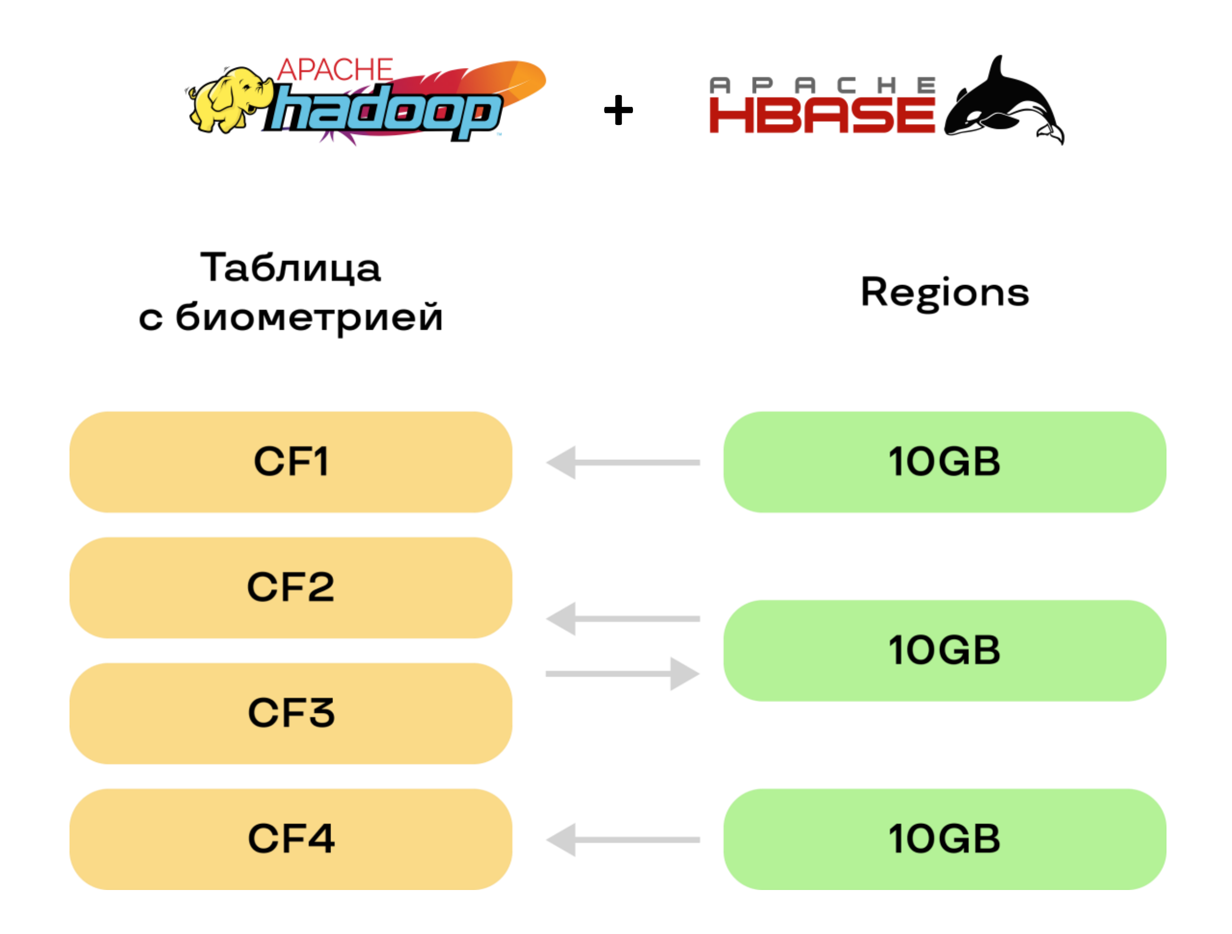






Передавать научились – давайте сохраним!

2. Размазывание нагрузки на кластер



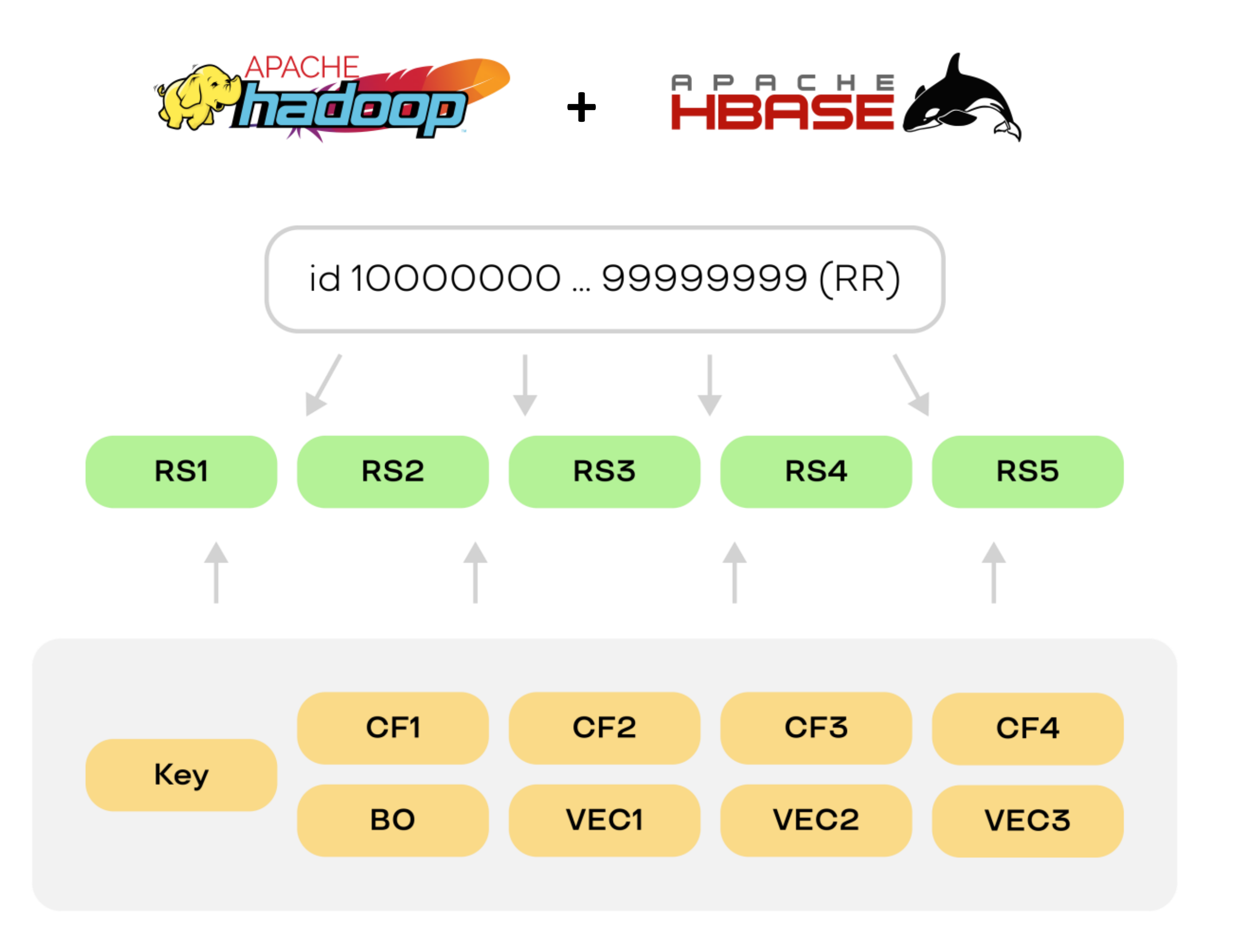






Передавать научились – давайте сохраним!

3. Хранение (Схема хранения биометрии по CF + запросы)









Этап 5

API

Модули ЕБС

Хранение

Транспорт

Работа с Liveness

Работа с биометрией

Измерение показателей нейросетей







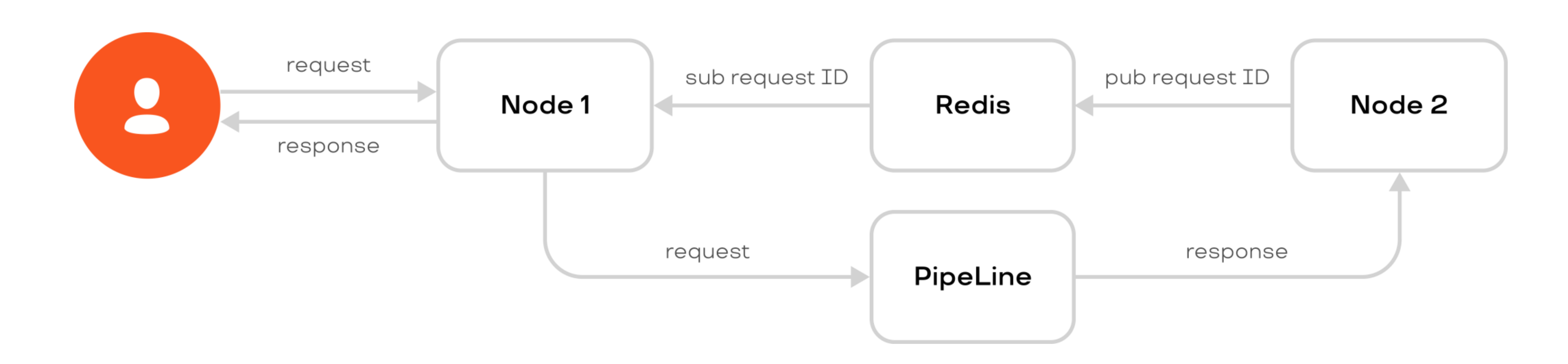
Мультимодальность, мульти-Liveness







Реализация синхронных “http”-запросов, снаружи в режиме НА с асинхронным взаимодействием внутри через Redis









PubSub для изменения настроек

Админ

Админка

Redis

Софт в OpenShift

docker

sub

pub





Zookeeper для кворума и почему от него отказались в своем ПО

Feature	ZK	OS
Centralized	+	+
Persistence	+	+
UI	–	+
APILess	–	+







Этап 6

API

Модули ЕБС

Хранение

Транспорт

Работа с Liveness

Работа с биометрией

Измерение показателей нейросетей







Внешнее взаимодействие

ПРОБЛЕМЫ

- Для Liveness много различных решений
- Возможность использования модальности зависит от устройств
- Много параметров в API – сложно (как дать клиенту простое API?)

РЕШЕНИЯ

- Сами реализовываем логику инструкций и проверки на стороне платформы.
- Поддерживаем различных клиентов и согласуем наборы действий в рамках протокола.
- Реализовываем сложное «внутри», сами делаем сложные интерфейсы.







Сложили картинку

API

Модули ЕБС

OPENSHIFT

Хранение

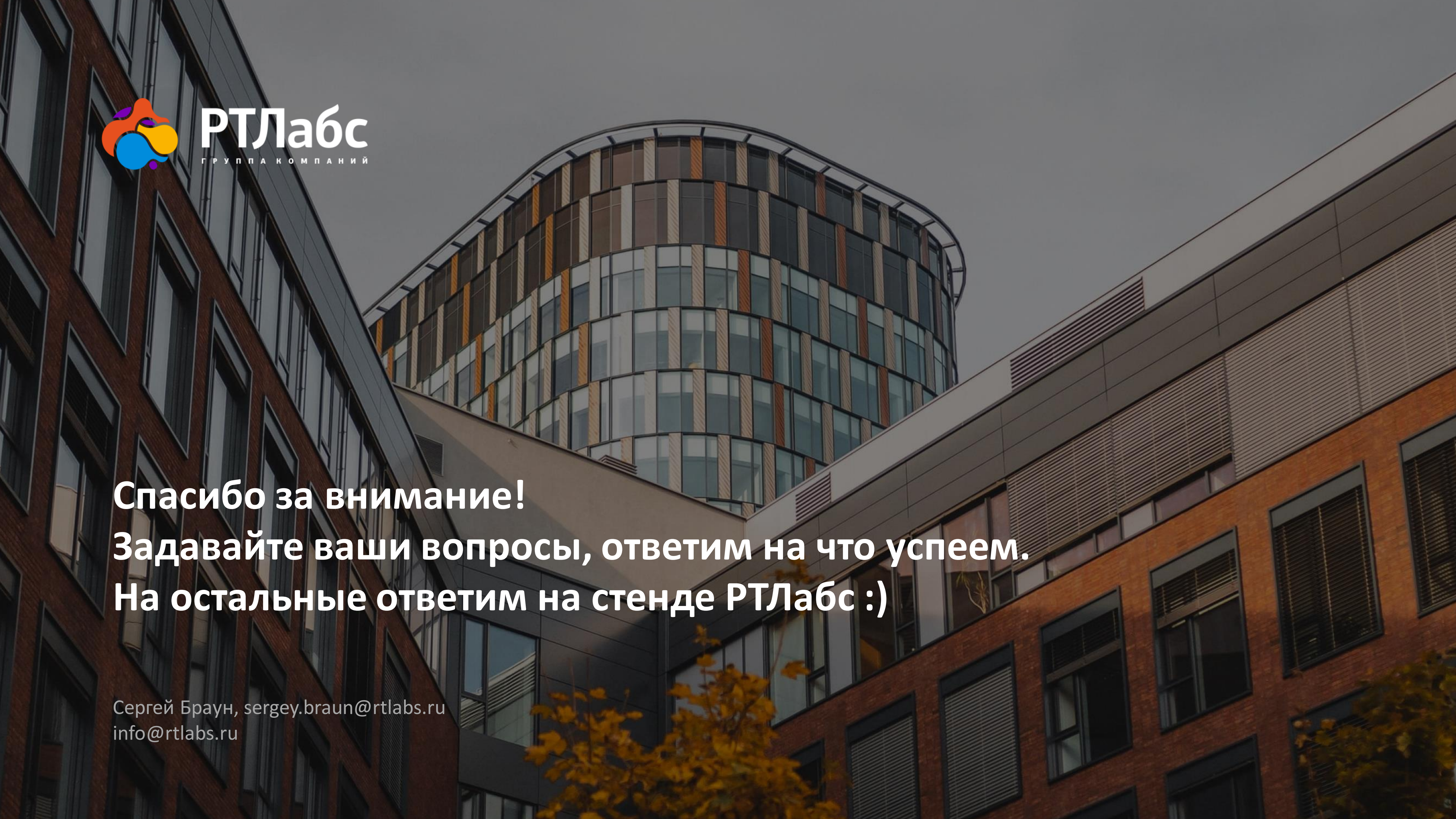
Транспорт

Работа с Liveness

Работа с биометрией

docker

Измерение показателей нейросетей


РТЛабс
ГРУППА КОМПАНИЙ
Спасибо за внимание!
Задавайте ваши вопросы, ответим на что успеем.
На остальные ответим на стенде РТЛабс :)
Сергей Браун, sergey.braun@rtlabs.ru
info@rtlabs.ru